

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-210196

(43)Date of publication of application : 11.08.1995

(51)Int.Cl.

G10L 7/04

H03M 7/30

H04N 7/24

(21)Application number : 06-339748

(71)Applicant : HEWLETT PACKARD CO <HP>

(22)Date of filing : 29.12.1994

(72)Inventor : KONSTANTINIDES KONSTANTINOS

(30)Priority

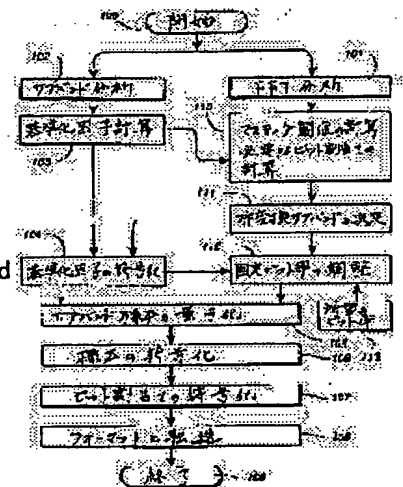
Priority number : 93 174747 Priority date : 29.12.1993 Priority country : US

(54) DEVICE AND METHOD FOR ENCODING/DECODING DIGITAL SIGNAL

(57)Abstract:

PURPOSE: To provide a method and device by which subbands can be filtrated at a high speed in encoding/decoding MPEG speech signals.

CONSTITUTION: An input sample is mapped to a frequency coefficient by using a subband analyzing filter (102) and a scaling factor is transmitted at every continuous output sample of each subband (103). In parallel with the transmission, an input speech signal is Fourier-transformed at a high speed (101) and a masking threshold is decided (110). Then a timbre signal component, non-timbre signal, and silence signal are calculated (111) and bit assignment is repeatedly calculated (112). Other bits are calculated for encoding each frame (106) and bit assignment information is transmitted (107). In addition, an encoded bit stream is outputted from an encoder for storage and arithmetic processing of addition, etc., (108).



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

BEST AVAILABLE COPY

* NOTICES *

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. **** shows the word which can not be translated.

3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] The phase of being an approach for encoding a sound signal and digitizing a voice source signal, Since a multiplier set is produced as a result, by carrying out subband filtering, said digitization voice source signal using a reverse dispersion cosine transformation operation The phase which maps said digitization voice source signal in a frequency domain from a time domain, Said method of having the phase of normalizing said multiplier by the selected scale factor, the phase which quantizes and carries out entropy code modulation of said multiplier, and the phase which carries out the multiplication of said multiplier since the encoded bit stream is produced.

[Claim 2] The approach according to claim 1 of encoding said sound signal according to MPEG voice specification.

[Claim 3] The approach according to claim 2 said mapping stage story has a 32-point reverse dispersion cosine transformation operation further.

[Claim 4] the data vector y by which said mapping stage story is further defined as $y''(k)$ is a degree type — " — (k —) — ($k = 0, 1, \text{ and } \dots 31$) — forming — [Equation 1]

$$y''(k) = \begin{cases} y(16) & k = 0 \\ y(k+16) + y(16-k) & k = 1, 2, \dots, 16 \\ y(k+16) - y(80-k) & k = 17, 18, \dots, 31 \end{cases}$$

Reverse dispersion cosine transformation of $y''(k)$ is performed as a degree type, and it is [Equation 2].

$$S(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2i+1)k\right] y''(k), \quad i = 0, 1, \dots, 31$$

The approach according to claim 3 of drawing output [of an MPEG encoder] $S(i)$ from the 32-point reverse dispersion cosine transformation operation of $y''(i)$ in a front type.

[Claim 5] The phase of unpacking a data frame including the digitization voice source signal which is an approach for decoding a sound signal and was encoded, Since an output sample set is produced as a result, by carrying out subband filtering, said digitization voice source signal using a discrete cosine transformation operation Said method of having the phase which reverse-maps said coding digitization voice source signal in a time domain from a frequency domain, and the phase which multiplexes said output sample since the decoded bit stream is produced.

[Claim 6] The approach according to claim 5 of decoding said sound signal according to MPEG voice specification.

[Claim 7] The approach according to claim 6 said reverse mapping stage story has a 32-point dispersion cosine transformation operation further.

[Claim 8] said — reverse — a mapping stage — a story — further — V — " — a degree — a type — it is — a passage — giving a definition — having — S — (k —) — dispersion — cosine transformation — V — " — carrying out — things — [Equation 3]

$$V''(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)i\right] S(k), \quad i = 0, 1, \dots, 31$$

asking for data vector $V(i)$ as follows — containing — the case of $i=0-15$ — $V(i) = V''(i+16)$, $V(32-i) = -$

$V(i) V(16) = 0$, and $V(48) = -V''(0)$ —

coming out — it is — i — = — one — 15 — a case — $V(48-i)$ — = — $-V''(-i)$ — and — $V(48+i)$ — = — $-V''(i)$ — it is — front — a formula — inside — MPEG — a decoder — an output — $V''(-i)$ — $S(-i)$ — 32 — a point — reverse — dispersion — cosine transformation — an operation — $V''(i)$ — from — drawing — being according to claim 7 — an approach .

[Claim 9] The A-D converter for being equipment for encoding a sound signal and digitizing a voice source signal, The subband analysis filter which uses the reverse dispersion cosine transformation operation for mapping said digitization voice source signal in a frequency domain from a time domain so that a multiplier set may be produced as a result, Said equipment which has the means for normalizing said multiplier by the selected scale factor, a means for quantizing and carrying out entropy code modulation of said multiplier, and a multiplexer for combining said multiplier since the encoded output bit stream is produced.

[Claim 10] Equipment according to claim 9 with which said sound signal is encoded according to MPEG voice specification.

[Claim 11] Equipment according to claim 10 which has a processor for said subband analysis filter to perform a 32-point reverse dispersion cosine transformation operation further.

[Claim 12] the data vector y by which said processor is further defined as $y''(k)$ is a degree type — " — $(-k)$ — $(-k = 0, 1, \text{ and } \dots \text{ the means for forming } 31)$, and [Equation 4]

$$y''(k) = \begin{cases} y(16) & k = 0 \\ y(k+16) + y(16-k) & k = 1, 2, \dots, 16 \\ y(k+16) - y(80-k) & k = 17, 18, \dots, 31 \end{cases}$$

It has a means for performing reverse dispersion cosine transformation of $y''(k)$ as a degree type, and is [Equation 5].

$$S(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2i+1)k\right] y''(k), \quad i = 0, 1, \dots, 31$$

Equipment according to claim 11 with which output [of an MPEG encoder] $S(i)$ is drawn from the 32-point reverse dispersion cosine transformation of $y''(i)$ in a front type.

[Claim 13] Said equipment which has the means for performing the unpack of a data frame including the digitization voice source signal which is a means for decoding a sound signal and was encoded; a subband analysis filter for reverse-mapping said digitization voice source signal in a time domain from a frequency domain using a discrete cosine-transformation operation so that an output sample set may produce as a result; and a multiplexer for combining said output sample since the decoded bit stream produces.

[Claim 14] Equipment according to claim 13 with which said sound signal is decoded according to MPEG voice specification.

[Claim 15] Equipment according to claim 14 with which said subband analysis filter contains the processor for performing a 32-point dispersion cosine transformation operation further.

[Claim 16] said — a processor — a means — further — V'' — a degree — a type — it is — a passage — giving a definition — having — $S(-i)$ — dispersion — cosine transformation — $V''(i)$ — carrying out — a sake — a means — [Equation 6]

$$V''(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)i\right] S(k), \quad i = 0, 1, \dots, 31$$

the means for asking for data vector $V(i)$ like a degree type — having — the case of $i=0-15$ — $V(i) = V''(i+16)$, $V(32-i) = -V''(i)$ $V(16) = 0$, and $V(48) = -V''(0)$ —

coming out — it is — i — = — one — 15 — a case — $V(48-i)$ — = — $-V''(-i)$ — and — $V(48+i)$ — = — $-V''(i)$ — it is — front — a formula — inside — MPEG — a decoder — an output — $V''(-i)$ — $S(-i)$ — 32 — a point — reverse — dispersion — cosine

transformation — an operation — V — " — (— i —) — from — drawing — having — being according to claim 15 — equipment .

[Translation done.]

*** NOTICES ***

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.*** shows the word which can not be translated.

3.In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to data coding / decryption. Furthermore, especially this invention relates to subband filtering in the case of coding of data, and a decryption (subband filtering).

[0002]

[Description of the Prior Art] Although the current storage makes it possible to store a huge quantity of information in the physical space of a critical mass, in order to perform a data compression, it is still desirable to use coding/decryption technique. In this case, stored data can be given in various suitable specification formats.

[0003] Decoding such encoded data needs for stored data to be restored by use of the various compression discharge technique. such restoration — count — enough like [it is intensive, and must be reliable and / in the case of being speech information and video information] to express stored data as real time — it must be carried out by being high-speed. in order [in addition,] to satisfy the requirement of a rate and accuracy — *** with superfluous decode processing — don't need expensive hardware.

[0004] A difference with the coding format big when satisfying two requirements of a rate and accuracy chosen may be brought about. There is various specification in the format used in order to encode a sound signal and a video signal for storage and retrieval. For example, Motion Picture Expert Group (MPEG) has established the specification for coding/decryption with a sound signal and a video signal. MPEG voice coding / decryption specification is incompressible Compact. High quality voice is brought about not using the bit rate of 706 kbit/s needed with the quality voice of Disc (trademark) but using a low bit rate called 128 kbit/s per voice channel. ISO/IEC 11172-3, "[Coding] of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s—Part3:Audio", ISO/IEC-JTC1/SC29, and 5/20/1993, H. Please refer to Musmann, The ISO Audio Coding Standard, IEEE GLOBECOM Conference, 511 – 517 pages, San Diego, and CA 1990.

[0005] By MPEG voice coding / decryption specification, since the bit stream into which it was compressed for storing, for example on the desirable media of arbitration, such as a compact disk and a tape, is produced, a digital sound signal is encoded. In MPEG voice coding, the non-standard coding technique is able to be used for a task like presumption of an audibility masking threshold, quantization, and/or normalization. This MPEG specification needs that it must be the coded signal which can produce the voice to which the decoder with which the encoded signal is based on the specification of

that specification fitted the application meant from that coded signal.

[0006] Although great efforts have been spent in order to improve the video decode technique and the display technique of an MPEG file, the efforts performed in order to improve the rate and effectiveness of MPEG voice coding are very slight an old place. As the result, voice decode formed 40% or more of the whole decode time amount at the beginning. Various improvements to MPEG voice decode are made, and decode of the voice part of one MPEG coding sequence has reached 33% of the whole decode time amount of the sequence. On the other hand, video coding reached 55% of the whole coding time amount, and synchronization has reached 12% of the whole decode time amount. use of hardware with high-speed special activation of MPEG specification — application of such a technique — superfluous — high — in most applications made cost, real time or when realizing coding and decode of real time in general, it is very important.

[0007] subband filtering — the inside of MPEG voice coding / decryption specification — most — count — it is an intensive operation, therefore is the field to which the serious interest is attracted in the improvement of coding/decryption engine performance about voice data. subband filtering — especially — count — one of the intensive operations is matrix operation. At present, matrix operation is performed by very troublesome count. That is, by MPEG specification about voice, $64 \times 2,048 \times 32$ "multiplication-accumulation" operation must be performed. Performing such an operation at the same time it satisfies the requirement of a rate and accuracy will add great complexity and costs to the design of the hardware which performs such a function.

[0008]

[Problem(s) to be Solved by the Invention] This invention offers the approach and equipment for high-speed subband filtering in MPEG voice coding / decryption.

[0009]

[Means for Solving the Problem] The sound signal digitized during coding/decryption processing is pretreated. This pretreated signal is sent to an encoder and matrix count in a voice coding subband analysis filter is performed within this encoder using a reverse dispersion cosine transformation function (inverse discrete cosinetransform function "IDCT"). This IDCT makes it possible to perform matrix count using 80 multiply operation and 209 add operations. The encoded data are stored.

[0010] During decode of storing data, after treatment of the storing bit stream is carried out before the decode with a voice decode subband composition filter. The matrix operation of a decode subband filter uses a discrete cosine transformation (discrete cosine transform (DCT)) operation.

[0011] This invention realizes remarkable improvement in a rate in MPEG voice coding / decryption processing. In MPEG decode, this invention brings about about 17% of improvement in the whole decode with which about 30% of improvement in voice decode, and voice decode and video decode were combined, without needing special hardware.

[0012]

[Example] This invention realizes more nearly high-speed coding/decryption for MPEG voice. MPEG voice coding / decryption specification makes it possible to reduce the sampling rate of a stereo sound signal from 2x706 kbit/s which is a sampling rate for compact disk stereos to 2x128 kbit/s. Such coding/a decryption improve the storage capacity of a storage like a compact disk greatly, for example, is increased to the storage capacity to 5 times of possible music storage capacity in current. If this invention is applied to MPEG voice coding / decryption, it turns out that about 17% of improvement in the whole decode with which about 30% of improvement in voice decode, and voice decode and video decode were combined is brought about without needing special hardware.

[0013] Next, MPEG voice coding / decryption specification is explained briefly. Drawing 1 is the schematic block diagram of the MPEG voice coder circuit 10. In this drawing, the input signal 11 which consists of pulse-code-modulation signals which have the sampling frequency of 48kHz and the sample size of 16 bits per sample is given to the single channel encoder 10. This signal is first mapped by the subband filter bank 12 from a time domain in a frequency domain. The multiplier obtained as the result is

normalized using the scale factor transmitted as typically secondary information. After that, entropy code modulation of the multiplier obtained in this way is quantized and carried out by quantization / coding network 14. The masking threshold of a quantization error is calculated based on the mental sound value given with the mental sound model (psychoacoustic model) 17 in order to control a quantization phase. The bit allocation to these multipliers is drawn from a masking function. This bit allocation information chooses one quantizer from 1 set of selectable quantizers. This bit allocation is transmitted as secondary information. In next, with an auxiliary data 13, the multiplication of the encoded signal is carried out by the frame pack circuit (frame packing circuit) 16, and it is generated for coding/decryption bit stream 18 in the output of an encoder 10.

[0014] Drawing 2 is the schematic block diagram of the MPEG voice decoder circuit 20. The coding bit stream 21 is given to the input of a decoder in this drawing. When an error correction operation is applied with an encoder, a bit stream unpack / decoder circuit 22 performs an error correction operation. It unpacks this bit stream data in order to restore each of the encoded information, and the auxiliary data 23 encoded by beforehand. The restoration circuit 24 restores the quantization version of a mapping sample set. The reverse mapping circuit 25 changes the mapped sample, and returns to the uniform pulse-code-modulation output signal 26 which is in agreement with the input signal over an encoder.

[0015] The above-mentioned explanation about an encoder and a decoder is peculiar to MPEG specification. Therefore, it is considered that it is handling within the limits of this contractor to materialize the above-mentioned various hardware functions. Therefore, this invention comes out enough and it is thought for a perfect understanding that still more detailed explanation of the hardware side of MPEG coding / decryption system is unnecessary. Although this invention mentioned on these specifications relates with MPEG coding / decryption specification and is explained, please understand that to be usable is considered by other coding/decryption applications and coding/decryption specification.

[0016] Drawing 3 is a flow chart which shows signal processing in the layers (layer) I and II in an MPEG voice coder. The sound signal by which pulse code modulation was carried out is given to an MPEG encoder (100). Since an input sample is mapped for a frequency multiplier, the subband analysis filter which has the subband of equal spacing of 32 pieces is used for MPEG voice specification (102). As a result of the division in the output of this subband filter, the output sampling rate of each filter is 1/32 of an input sampling rate. These subband filters are finite impulse response filters of the length 512, and are adaptation-ized by performing anti-aliasing.

[0017] A scale factor is transmitted for every 12 continuous output samples of each subband (103). This scale factor is expressed with 6 bits in accordance with the maximum absolute value of 12 samples. In the case of entropy code modulation, in order to transmit a scale factor, the data transfer rate of 8 kbit/s is needed.

[0018] In order are parallel to a subband filtering operation and to obtain high-resolution presumption of a power spectrum, the 1024-point fast Fourier transform of an input sound signal is calculated (101). A power spectrum is analyzed in order to determine a masking threshold (110). Therefore, in order that MPEG specification may calculate the masking threshold and signal pair mask ratio (signal-to-mask ratio) in each subband (SMR), each of a tone signal component (tonal signal component), a non-tone signal (non-tonal signal component), and a silence signal (silent signal component) is taken into count (111). Bit allocation is a repetitive procedure (112 113). In each repetitive step, it follows with the number of quantization level, and signal to quantization noise ratio (signal-to-quantizing noise coderatio) (SNR) is a mask pair quantization ratio (mask-to-quantizing ratio) (MNR).

MNR = Only the decided increment is increased in the subband whose SNR-SMR is min (105). Other bits for being able to come, simultaneously encoding the frame are calculated (106). The above-mentioned recursive procedure is repeated as long as there are the remaining bits. Bit allocation information (107) is transmitted every 24ms, therefore the data transfer rate of 3.5 kbit/s is needed

(108). The encoded bit stream is outputted by the encoder for storing, addition processing, etc.

[0019] Drawing 4 is a flow chart which shows signal processing in the layers I and II in an MPEG voice decoder. Bit allocation (200 201) of an input bit stream is decoded (202). After that, various scale factors are also decoded (203) and a sample is re-quantized (204). The encoded signal is decoded with a synthetic subband filter (205), and the decoded pulse-code-modulation signal is outputted for still more nearly another processing and/or real-time playback (206 207).

[0020] Below, the desirable example of this invention is explained to a detail.

[0021] (MPEG encoder) Drawing 5 is a flow chart which shows actuation of the analysis subband filter for an MPEG voice coder (as being shown to drawing 3 by the reference number 102). This drawing is main filter actuation (the shift only of the 64 figures of the data is carried out) (301, 302, 303), i.e., a shift operation, and a window operation (since an output sample is produced). The multiplication of the data by which the shift was carried out is carried out by the windowing function, it is added after that (303), and the matrix operation (64x32-line column vector multiplication is performed) (305) for generating the decoded output signal (306 307) is shown.

[0022] It defines as follows and the matrix operation (305) in an MPEG voice coding subband analysis filter is [0023].

[Equation 7]

$$S(i) = \sum_{k=0}^{31} \cos \left[\frac{\pi}{64} (2i+1)(k-16) \right] y(k) \quad , i = 0, 1, \dots, 31 \quad (1)$$

[0024] In a front type, after a window operation, a shift operation, and an add operation are performed as it defined as the above-mentioned specification, 64 y (k) samples are drawn from 32 input voice samples x (i) (302, 303, 304).

[0025] When it is k= 0, and (1, 2, ..., 63), the matrix operation in the subband analysis filter of this invention is as follows. [input y (k), and]

1. When defining as y' (k) is a degree type, k= 0, and (1, 2, ..., 63) are formed, and it is [0026]. [data vector y' (k), and]

[Equation 8]

$$y'(k) = \begin{cases} y(k+16) & k = 0, 1, \dots, 47 \\ -y(k-48) & k = 48, 49, \dots, 63 \end{cases} \quad (2)$$

[0027] At this time, it is [0028].

[Equation 9]

$$S(i) = \sum_{k=0}^{31} \cos \left[\frac{\pi}{64} (2i+1)k \right] y'(k) \quad , i = 0, 1, \dots, 31 \quad (3)$$

[0029] 2. When it is as y' (k) having been defined above (2), a data vector is formed as follows.

$$y''(0) = y'(0)$$

$$y''(k) = y'(k) - y'(64-k) \quad k=1, \dots, 31 \quad (4)$$

At this time, it is [0030].

[Equation 10]

$$S(i) = \sum_{k=0}^{31} \cos \left[\frac{\pi}{64} (2i+1)k \right] y''(k) \quad , i = 0, 1, \dots, 31 \quad (5)$$

[0031] a data sequence x (i) — (— i= — X (i) expresses 1, 2, ..., the discrete cosine transformation of N-1). In this case, it defines as follows and X (i) is [0032].

[Equation 11]

$$X(i) = \frac{2}{N} a(i) \sum_{k=0}^{N-1} x(k) \cos \left[\frac{\pi}{2N} (2k+1)i \right] \quad , i = 0, 1, \dots, N-1 \quad (6)$$

[0033] It is expressed with a degree type and IDCT of X (i) is [0034].

[Equation 12]

$$x(i) = \sum_{k=0}^{N-1} e(k)X(k) \cos\left[\frac{\pi}{2N}(2i+1)k\right], \quad i = 0, 1, \dots, N-1 \quad (7)$$

[0035] In a front type, when it is $i = 0$, it is $e(i) = 1/\sqrt{2}$, and when other, it is $e(i) = 1$.

[0036] An equation (5) needs 32×32 multiplication accumulation as shown above. As compared with an equation (7), $S(i)$ corresponds the above-mentioned equation (5) to the 32-point reverse dispersion cosine transformation ("IDCT") of $y''(k)$ (within the limits of one criteria factor). Therefore, with the subband filter of this invention, output [of MPEG encoder matrix operation] $S(i)$ is drawn from 32-point IDCT of $y''(i)$.

[0037] By combining the definition of the above-mentioned equation (2), and the definition of an equation (4), the direct relation between $y''(k)$ and $y(k)$ is able to be drawn as follows.

[0038]

[Equation 13]

$$y''(k) = \begin{cases} y(16) & k = 0 \\ y(k+16) + y(16-k) & k = 1, 2, \dots, 16 \\ y(k+16) - y(80-k) & k = 17, 18, \dots, 31 \end{cases} \quad (8)$$

[0039] (MPEG decoder) Drawing 6 is a flow chart which shows actuation of the synthetic subband filter for an MPEG voice decoder. MPEG voice decode processing (layers I and II) includes a synchronization function, a grouping discharge (degrouing) function, a quantization discharge (dequantization) function, a normalization discharge (denormalization) function, and a subband composition function (as being shown to drawing 4 by the reference number 205) as shown in drawing 4. At drawing 6, 32 subband samples are inputted into a decoder (400 401), and matrix operation is performed after a shift operation (402) (403 404). After a window operation (405), 32 decoded samples are calculated (406) and the restored pulse-code-modulation signal is outputted (407 408).

[0040] The matrix operation in a subband composition filter (as being shown to drawing 6 by the reference number 403) is defined by MPEG voice decode by the degree type.

[0041]

[Equation 14]

$$V(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)((i+16))\right] S(k), \quad i = 0, 1, \dots, 63 \quad (9)$$

[0042] Troublesome count of the above-mentioned equation (9) needs a $32 \times 64 = 2048$ time multiplication accumulation operation. The performance analysis about MPEG voice decode shows that matrix operation occupies about 40% of the whole decode time amount. With the subband composition filter of this invention, matrix operation makes decode possible by performing an equation (9) using 32-point DCT.

[0043] Matrix operation in the subband composition filter of this invention is performed as follows.

[0044] When $V(i)$ is given from an equation (9), $V'(i)$ is defined as follows.

[0045]

[Equation 15]

$$V'(i) = \begin{cases} V(i+48) & i = 0, 1, \dots, 15 \\ V(i-16) & i = 16, 17, \dots, 63 \end{cases} \quad (10)$$

[0046] (10) to the above-mentioned equation (9) and a degree type are obtained.

[0047]

[Equation 16]

$$V'(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)(i+64)\right] S(k), \quad i = 0, 1, \dots, 15 \quad (11)$$

$$V'(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)i\right] S(k), \quad i = 16, 17, \dots, 63 \quad (12)$$

$$V'(32) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{2}(2k+1)\right] S(k) = 0 \quad (13)$$

[0048] About $i = 17$, and $18, \dots, 31$, when defining as V' is the above, it is $V'(32+i) = -V'(32-i)$ about $i = 1$, and $2, \dots, 16$, and is $V'(32+i) = V'(32-i)$.

[0049] V'' — $i = 0$, and $1, \dots, 15$ — being related — V'' — a definition is given as $(i) = -V'(i)$ — having — $i = 16$, and $17, \dots, 31$ — being related — V'' — the case where it defines as $(i) = V'(i)$ — [0050]

[Equation 17]

$$V''(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)i\right] S(k), \quad i = 0, 1, \dots, 31 \quad (14)$$

[0051] From the above-mentioned equation (6) and (14), $V''(i)$ is in agreement with the 32-point dispersion cosine transformation of $S(i)$ (within the limits of one scale factor). Vector V which should be searched for (i) It is obtained from $V''(i)$ as follows.

In the case of $i=0-15$ $V(i) = V''(i+16)$ $V(32-i) = -V(i)$ $V(16) = 0$ (15)

Come out, and it is and, in the case of $i=1-15$, is. $V(48-i) = -V''(i)$ $V(48+i) = -V''(i)$ $V(48) = -V''(0)$ (16)

[0052] Therefore, the output V of the above-mentioned matrix operation equation (9) in MPEG voice decode is obtained after the after treatment of DCT conversion of S .

[0053] Drawing 7 is the flow chart of the matrix operation of the DCT base in the synthetic subband filter for the MPEG voice decoder by this invention. Therefore, 32 samples are inputted into an encoder (500 510), and V'' is calculated from 32-point DCT conversion of S (501 510). From step 2 on the left-hand side of drawing, the right-hand side of this drawing shows that being combined with the form of one step (512) is possible so that V may be able to calculate step 4 (502, 503, 504) directly from V'' . After that, the decoded pulse-code-modulation signal is outputted by the decoder (505 513).

[0054] This invention enables more effective MPEG voice coding / decryption by using IDCT for coding for decode, using DCT in the one side face by enabling more nearly high-speed activation of the matrix operation under subband filtering. In addition, this invention decreases the memory requirement of a system by calculating with the simplified cosine multiplier set.

[0055] Many technique for performing the operation of IDCT and DCT is known (please refer to KR.Rao, P.Yip, Discrete Cosine Transform, Academic Press, and 1990). For example, the high-speed DCT operation of Lee is repetitive, it is possible to be applied to 32-point DCT (or IDCT), and 80 multiply operation and 209 add operations are performed in total as the result (B. G.Lee, A New Algorithm To Compute The Discrete Cosine Transform, IEEE Trans.on ASSP, Vol.ASSP-32, No.6, pp.1243-1245, Dec.1984). This shows saving on remarkable count as compared with troublesome count which is used with the conventional technique.

[0056] Since the algorithm of Lee is efficient and an understanding, and coding/decryption are easy for it, it is used in the desirable example of this invention. The activation of the 32-point high speed DCT based on the algorithm of Lee is explained below.

[0057] DCT data sequence $x(i) =$ — if $0, 1, \dots, N-1$ are expressed with $X(i)$, in the case of $N= 32$, the equation for a high speed DCT will be given by the degree type.

[0058]

[Equation 18]

$$G(n) = X(2n), \quad n = 0, 1, \dots, 15 \quad (17)$$

$$H(n) = X(2n+1) + X(2n-1), \quad n = 0, 1, \dots, 15 \quad (18)$$

$$x(k) = g(k) + \frac{1}{2c_{2k+1}^{2k+1}} h(k) \quad (19)$$

$$x(31-k) = g(k) - \frac{1}{2c_{2k+1}^{2k+1}} h(k), \quad k = 0, 1, \dots, 15 \quad (20)$$

[0059] However, [0060]

[Equation 19]

$$c_{2N}^{(2k+1)n} = \cos(\pi(2k+1)n/2N) \quad (21)$$

[0061] Drawing 8 A and drawing 8 B are the data flow charts about the 32-point high-speed DCT operation by the algorithm of Lee performed by this invention. From an equation (17) to an equation (20) forms the 1st phase and culmination of the flow chart shown in drawing 8 A and drawing 8 B (by arrow head of the reverse sense). The values H, G, g, and h in both drawings are repetitively calculated using the same equation for a 16-point DCT operation, an eight-point DCT operation, a four-point DCT operation, and a two-point DCT operation. In order to calculate IDCT for MPEG coding, the processing direction shown in drawing is reversed. At drawing 8 A and drawing 8 B, it is [0062].

[Equation 20]

$$C(2N, k) = 1 / \left(2c_{2N}^k \right) \quad (22)$$

[0063] Although this specification has explained this invention with reference to a desirable example, this contractor will understand that this invention can be applied to applications other than the application explained on these specifications, without deviating from the thought of this invention, and the range. Therefore, this invention is limited by only the attachment claim.

[0064] As mentioned above, although the example of this invention was explained in full detail, this inventions are hereafter enumerated for every embodiment.

(1). ~~As mentioned above, although the example of this invention was explained in full detail, this inventions are hereafter enumerated for every embodiment.~~

[0065] ~~As mentioned above, although the example of this invention was explained in full detail, this inventions are hereafter enumerated for every embodiment.~~

[Effect] As mentioned above, if this invention is used, the improvement in a rate in MPEG voice coding / decryption processing is realizable. In MPEG decode, this invention brings about about 17% of improvement in the whole decode with which about 30% of improvement in voice decode, and voice decode and video decode were combined, without needing special hardware.

[Translation done.]

* NOTICES *

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.**** shows the word which can not be translated.

3.In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the schematic block diagram of an MPEG voice coder.

[Drawing 2] It is the schematic block diagram of an MPEG voice decoder.

[Drawing 3] It is the flow chart which shows signal processing in Layer I and Layer II of an MPEG voice coder.

[Drawing 4] It is the flow chart which shows signal processing in Layer I and Layer II of an MPEG voice decoder.

[Drawing 5] It is the flow chart which shows actuation of the analysis subband filter for MPEG voice coders.

[Drawing 6] It is the flow chart which shows actuation of the analysis subband filter for MPEG voice decoders.

[Drawing 7] It is the flow chart of the matrix operation of the DCT base in the synthetic subband for the voice MPEG decoder by this invention.

[Drawing 8 A] They are the 32-point high speed DCT by this invention, and a data flow chart about an IDCT operation.

[Drawing 8 B] They are the 32-point high speed DCT by this invention, and a data flow chart about an IDCT operation.

[Description of Notations]

10: MPEG voice coder

11: Input signal

12: Subband filter

14: Quantization/coding network

13: Auxiliary data

17: Mental sound model

20: MPEG voice decoder

21: Coding bit stream

22: A bit stream unpack / decoder circuit

23: Auxiliary data

24: Restoration circuit

25: Reverse mapping circuit

26: Output signal

100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113: Block

[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-210196

(43) 公開日 平成7年(1995) 8月11日

(51) Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 7/04	G			
H 0 3 M 7/30	A	8842-5 J		
H 0 4 N 7/24				
			H 0 4 N 7/ 13	Z

審査請求 未請求 請求項の数16 F D (全 10 頁)

(21) 出願番号 特願平6-339748

(22) 出願日 平成6年(1994)12月29日

(31) 優先権主張番号 1 7 4, 7 4 7

(32) 優先日 1993年12月29日

(33) 優先権主張国 米国 (US)

(71) 出願人 590000400

ヒューレット・パッカード・カンパニー
アメリカ合衆国カリフォルニア州パロアル
ト ハノーバー・ストリート 3000

(72) 発明者 コンスタンティノス・コンスタンティニデ
ス

アメリカ合衆国カリフォルニア州サンノ
ゼ、ジャコブ・アベニュー 1508

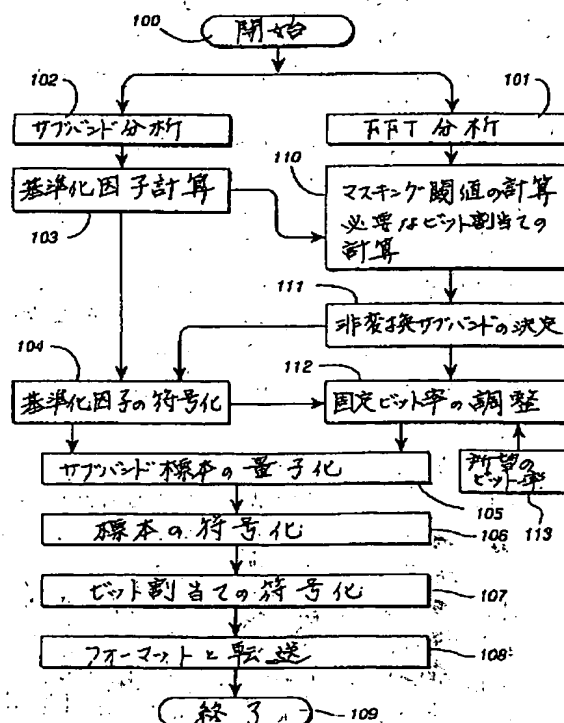
(74) 代理人 弁理士 上野 英夫

(54) 【発明の名称】 デジタル信号符号化/復号化装置及び方法

(57) 【要約】

【目的】 M P E G 音声符号化/複合化における高速サブ
バンドフィルタリングのための方法と装置を提供する。

【構成】 サブバンド分析フィルタを使用して、入力標本
を周波数係数にマッピングし (1 0 2)、各サブバンド
の連続した出力標本毎に基準化因子が伝送される (1 0
3)。これと平行して、入力音声信号の高速フーリエ変
換を行い (1 0 1)、マスキング閾値を決定する (1 1
0)。次に、音色信号成分、非音色信号、沈黙信号を計
算し (1 1 1)、ビット割当てを反復的に計算する (1
1 2、1 1 3)。各フレームを符号化するため他のビット
が計算され (1 0 6)、ビット割当て情報 (1 0 7)
が伝送され、符号化されたビットストリームが格納と加
算処理等のために符号器によって出力される (1 0
8)。



(2)

【特許請求の範囲】

【請求項1】 音声信号を符号化するための方法であつて、

音声ソース信号をディジタル化する段階と、

係数セットを結果的に生じさせるために逆離散余弦変換演算を使用して前記ディジタル化音声ソース信号をサブバンドフィルタリングすることによって、前記ディジタル化音声ソース信号を時間領域から周波数領域にマッピングする段階と、

選択された基準化因子によって前記係数を正規化する段階と、

前記係数を量子化しエントロピ符号化する段階と、

符号化されたビットストリームを生じさせるために前記係数を乗算する段階とを有する前記方法。

【請求項2】 前記音声信号をMPEG音声規格に従って符号化する請求項1に記載の方法。

【請求項3】 前記マッピング段階が更に、32点逆離散余弦変換演算を有する請求項2に記載の方法。

【請求項4】 前記マッピング段階が更に、

$y''(k)$ が次式の通りに定義される、データベクトル $y''(k)$ ($k=0, 1, \dots, 31$) を形成し、

【数1】

$$y''(k) = \begin{cases} y(16) & k=0 \\ y(k+16) + y(16-k) & k=1, 2, \dots, 16 \\ y(k+16) - y(80-k) & k=17, 18, \dots, 31 \end{cases}$$

次式の通りに $y''(k)$ の逆離散余弦変換を行い、

【数2】

$$S(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2i+1)k\right] y''(k), \quad i=0, 1, \dots, 31$$

前式中で、MPEG符号器の出力 $S(i)$ を、 y''

(i) の32点逆離散余弦変換演算から導き出す請求項3に記載の方法。

【請求項5】 音声信号を復号するための方法であつて、

符号化されたディジタル化音声ソース信号を含むデータフレームをアンパックする段階と、

出力標本セットを結果的に生じさせるために離散余弦変換演算を使用して前記ディジタル化音声ソース信号をサブバンドフィルタリングすることによって、前記符号化ディジタル化音声ソース信号を周波数領域から時間領域に逆マッピングする段階と、

復号されたビットストリームを生じさせるために前記出力標本を多重化する段階とを有する前記方法。

【請求項6】 前記音声信号をMPEG音声規格に従って復号する請求項5に記載の方法。

【請求項7】 前記逆マッピング段階が更に、32点離散余弦変換演算を有する請求項6に記載の方法。

【請求項8】 前記逆マッピング段階が更に、

V'' が次式の通りに定義される、 $S(k)$ の離散余弦変

換 V'' を行うことと、

【数3】

$$V''(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)i\right] S(k), \quad i=0, 1, \dots, 31$$

次のようにデータベクトル $V(i)$ を求めることとを含み、

$i=0 \sim 15$ の場合に、

$V(i) = V''(i+16)$ 、 $V(32-i) = -V(i)$ 、

$V(16) = 0$ 、 $V(48) = -V''(0)$

であり、 $i=1 \sim 15$ の場合に、

$V(48-i) = -V''(i)$ 且つ $V(48+i) = -V''(i)$ であり、

前式中で、MPEG復号器の出力 $V(i)$ を、 $S(i)$ の32点逆離散余弦変換演算 $V''(i)$ から導き出す請求項7に記載の方法。

【請求項9】 音声信号を符号化するための装置であつて、

音声ソース信号をディジタル化するためのAD変換器と、

係数セットを結果的に生じさせるように前記ディジタル化音声ソース信号を時間領域から周波数領域にマッピングするための逆離散余弦変換演算を使用するサブバンド分析フィルタと、

選択された基準化因子によって前記係数を正規化するための手段と、

前記係数を量子化しエントロピ符号化するための手段と、

符号化された出力ビットストリームを生じさせるために前記係数を組み合わせるためのマルチプレクサとを有する前記装置。

【請求項10】 前記音声信号がMPEG音声規格に従って符号化される請求項9に記載の装置。

【請求項11】 前記サブバンド分析フィルタが更に、32点逆離散余弦変換演算を行うためのプロセッサを有する請求項10に記載の装置。

【請求項12】 前記プロセッサが更に、

$y''(k)$ が次式の通りに定義される、データベクトル $y''(k)$ ($k=0, 1, \dots, 31$) を形成するための手段と、

【数4】

$$y''(k) = \begin{cases} y(16) & k=0 \\ y(k+16) + y(16-k) & k=1, 2, \dots, 16 \\ y(k+16) - y(80-k) & k=17, 18, \dots, 31 \end{cases}$$

次式の通りに $y''(k)$ の逆離散余弦変換を行うための手段とを有し、

【数5】

$$S(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2i+1)k\right] y''(k), \quad i=0, 1, \dots, 31$$

(3)

3

前式中で、MPEG符号器の出力 $S(i)$ が、 y''

(i) の32点逆離散余弦変換から導き出される請求項11に記載の装置。

【請求項13】 音声信号を復号するための手段であって、

符号化されたデジタル化音声ソース信号を含むデータフレームのアンパックを行うための手段と、

出力標本セットを結果的に生じさせるように離散余弦変換演算を使用して前記デジタル化音声ソース信号を周波数領域から時間領域に逆マッピングするためのサブバンド分析フィルタと、

復号されたビットストリームを生じさせるために前記出力標本を組み合わせるためのマルチプレксаとを有する前記装置。

【請求項14】 前記音声信号がMPEG音声規格に従って復号される請求項13に記載の装置。

【請求項15】 前記サブバンド分析フィルタが更に、32点離散余弦変換演算を行うためのプロセッサを含む請求項14に記載の装置。

【請求項16】 前記プロセッサ手段が更に、 V'' が次式の通りに定義される、 $S(i)$ の離散余弦変換 V'' を行うための手段と、

【数6】

$$V''(i) = \sum_{k=0}^{31} \cos \left[\frac{\pi}{64} (2k+1)i \right] S(k), \quad i = 0, 1, \dots, 31$$

次式のようにデータベクトル $V(i)$ を求めるための手段とを有し、

$i = 0 \sim 15$ の場合に、

$V(i) = V''(i+16)$ 、 $V(32-i) = -V(i)$ 、

$V(16) = 0$ 、 $V(48) = -V''(0)$

であり、 $i = 1 \sim 15$ の場合に、

$V(48-i) = -V''(i)$ 且つ $V(48+i) = -V''(i)$ であり前式中で、MPEG復号器の出力 $V(i)$ が、 $S(i)$ の32点逆離散余弦変換演算 $V''(i)$ から導き出される請求項15に記載の装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、データ符号化/復号化に係わる。更に特に、本発明は、データの符号化と復号化との際のサブバンドフィルタリング(subband filtering)に係わる。

【0002】

【従来の技術】最少量の物理的空間内に膨大な量の情報が格納されることを現在の記憶媒体が可能にしているにも係わらず、データ圧縮を行うために符号化/復号化技術を使用することが依然として望ましい。この場合、記憶データは様々な適切な規格フォーマットで与えられることが可能である。

4

【0003】こうした符号化されたデータを復号することは、様々な圧縮解除手法の使用によって記憶データが復元されることを必要とする。こうした復元は、計算集約的であり、信頼性が高くなければならず、且つ、例えば音声情報とビデオ情報の場合のように、記憶データをリアルタイムで表示するのに十分な高速度で行われなければならない。これに加えて、速度と正確度との必要条件を満たすために、復号処理が過剰な又は高価なハードウェアを必要としてはならない。

【0004】選択される符号化フォーマットが、速度と正確度という2つの要件を満たす上で、大きな差異をもたらす可能性がある。記憶と検索とのために音声信号とビデオ信号とを符号化するために使用されるフォーマットには、様々な規格がある。例えば、Motion Picture Expert Group(MPEG)は、音声信号とビデオ信号との符号化/復号化のための規格を確立している。MPEG音声符号化/復号化規格は、非圧縮のCompact Disc(商標)の品質音声で必要とされる706kbit/sのビット伝送速度ではなく、音声チャンネル当たり128kbit/sという低いビット伝送速度を使用して高品質音声をもたらす。ISO/IEC 11172-3、

「Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s -Part3:Audio」、ISO/IEC JTC 1/SC29,5/20/1993と、H. M. smann, The ISO Audio Coding Standard, IEEE GLOBECOM Conference, 511-517ページ, San Diego, CA 1990とを参照されたい。

【0005】MPEG音声符号化/復号化規格では、例えばコンパクトディスクやテープ等のような任意の望ましいメディア上に格納するための圧縮されたビットストリームを生じさせるために、デジタル音声信号が符号化される。MPEG音声符号化においては、非標準的な符号化手法が、聴感マスキング閾値の推定、量子化、及び/又は基準化のようなタスクのために使用されることが可能である。このMPEG規格は、符号化された信号が、その規格の仕様に準拠する復号器がその符号化信号から意図された用途に適した音声を生じさせることが可能であるような符号化信号でなければならないことを必要とする。

【0006】MPEGファイルのビデオ復号手法と表示手法とを改善するためには多大な努力が費やされてきたが、MPEG音声符号化の速度と効率とを改善するために行われた努力は、これまでのところ非常に僅かであるにすぎない。その結果として、音声復号は、当初には復号時間全体の40%以上を占めていた。MPEG音声復号に対する様々な改善が行われており、1つのMPEG符号化シーケンスの音声部分の復号は、そのシーケンスの復号時間全体の3.3%に達している。一方、ビデオ符号化は符号化時間全体の55%に達し、同期化は復号時間全体の12%に達している。MPEG規格の高速実行

(4)

5

は、特殊なハードウェアの使用がそうした技術の適用を過剰に高コストなものにする用途の殆どにおいて、リアルタイムの又は概ねリアルタイムの符号化と復号とを実現する上で非常に重要である。

【0007】サブバンドフィルタリングは、MPEG音声符号化／復号化規格の中で最も計算集約的な演算であり、従って、音声データに関する符号化／復号化性能の改善において重大な関心を集めている領域である。サブバンドフィルタリングの特に計算集約的な演算の1つは、行列演算である。現時点では、行列演算は非常に面倒な計算によって行われる。即ち、音声に関するMPEG規格では、 $32 \times 64 = 2,048$ 回の「乗算－累算」演算が行われなければならない。速度と正確度との必要条件を満たすと同時にこうした演算を行うことは、こうした関数を実行するハードウェアの設計に多大な複雑性と費用とを付加することになる。

【0008】

【発明が解決しようとする課題】本発明は、MPEG音声符号化／復号化における高速サブバンドフィルタリングのための方法と装置とを提供する。

【0009】

【課題を解決するための手段】符号化／復号化処理中に、ディジタル化された音声信号が前処理される。この前処理された信号は符号器に送られ、この符号器内では音声符号化サブバンド分析フィルタにおける行列計算が、逆離散余弦変換関数(inverse discrete cosinetran sform function(「IDCT」))を使用して行われる。このIDCTは、80回の乗算演算と209回の加算演算とを使用して行列計算が行われることを可能にする。符号化されたデータは格納される。

【0010】格納データの復号中に、音声復号サブバンド合成フィルタによる復号の前に格納ビットストリームが後処理される。復号サブバンドフィルタの行列演算は、離散余弦変換(discrete cosine transform(DCT))演算を使用する。

【0011】本発明は、MPEG音声符号化／復号化処理における著しい速度向上を実現する。MPEG復号では、本発明は、特別なハードウェアを必要とせずに、音声復号における約30%の改善と、音声復号とビデオ復号とが組み合わされた復号全体における約17%の改善とをもたらす。

【0012】

【実施例】本発明は、MPEG音声のためのより高速な符号化／復号化を実現する。MPEG音声符号化／復号化規格は、ステレオ音声信号の標本抽出率を、コンパクトディスクステレオ用の標本抽出率である $2 \times 706 \text{ kb/s}$ から $2 \times 128 \text{ kb/s}$ に低減させることを可能にする。こうした符号化／復号化は、コンパクトディスクのような記憶媒体の記憶容量を大きく改善し、例えば現在において可能な音楽記憶容量の5倍まで

6

の記憶容量に増大させる。本発明をMPEG音声符号化／復号化に適用すると、特別なハードウェアを必要とせずに、音声復号における約30%の改善と、音声復号とビデオ復号とが組み合わされた復号全体における約17%の改善とをもたらすことがわかっている。

【0013】次に、MPEG音声符号化／復号化規格を簡単に説明する。図1は、MPEG音声符号器回路10の略ブロック図である。この図では、48kHzの標本化周波数と標本1つ当たり16ビットの標本サイズとを有するパルス符号変調信号から構成される入力信号11が、単一チャンネル符号器10に与えられる。この信号は、最初に、サブバンドフィルタバンク12によって時間領域から周波数領域にマッピングされる。その結果として得られる係数が、典型的には副次的な情報として伝送される基準化因子を使用して正規化される。その後で、こうして得られた係数が、量子化／符号化回路14によって量子化されエントロピ符号化される。量子化誤差のマスキング閾値が、量子化段階を制御するため心理音響モデル(ps psychoacoustic model)17によって与えられる心理音響値に基づいて計算される。これらの係数に対するビット割当てが、マスキング関数から導出される。このビット割当て情報は、1組の選択可能な量子化器の中から1個の量子化器を選択する。このビット割当ては副次的な情報として伝送される。この後で、符号化された信号は、補助データ13と共に、フレームパック回路(frame packing circuit)16によって乗算され、符号化／復号化ビットストリーム18が符号器10の出力において発生させられる。

【0014】図2は、MPEG音声復号器回路20の略ブロック図である。この図では、符号化ビットストリーム21が復号器の入力に与えられる。誤差訂正演算が符号器で適用された場合には、ビットストリームアンパック／復号回路22が誤差訂正演算を行う。このビットストリームデータは、符号化された情報の各々と事前に符号化された補助データ23とを復元するためにアンパックされる。復元回路24は、マッピング標本セットの量子化バージョンを復元する。逆マッピング回路25が、マッピングされた標本を変換し、符号器に対する入力信号に一致する一様なパルス符号変調出力信号26に戻す。

【0015】符号器と復号器とに関する上記の説明は、MPEG規格に特有のものである。従って、上記の様々なハードウェア機能を具体化することは、当業者の取り扱い範囲内であると見なされる。従って、MPEG符号化／復号化システムのハードウェア面の更に詳細な説明は、本発明の十分で完全な理解にとって不必要であると考えられる。本明細書で言及される本発明が、MPEG符号化／復号化規格に関連付けて説明されているが、他の符号化／復号化用途と符号化／復号化規格とに使用可能であると考えられるということを理解されたい。

(5)

7

【0016】図3は、MPEG音声符号器でのレイヤ(I、II)における信号処理を示すフローチャートである。パルス符号変調された音声信号がMPEG符号器に与えられる(100)。MPEG音声規格は、入力標本を周波数係数にマッピングするために、32個の均等間隔のサブバンドを有するサブバンド分析フィルタを使用する(102)。このサブバンドフィルタの出力における分割の結果として、各フィルタの出力標本抽出率は、入力標本抽出率の1/32である。これらのサブバンドフィルタは、長さ512の有限インパルス応答フィルタであり、エイリアス除去を行うことに適合化されている。

【0017】基準化因子が、各サブバンドの12個の連続した出力標本毎に伝送される(103)。この基準化因子は、12個の標本の最大絶対値に一致し、6ビットで表される。エントロピ符号化の場合には、基準化因子を伝送するために8kbit/sのデータ転送速度が必要とされる。

【0018】サブバンドフィルタリング演算と平行して、パワースペクトルの高分解能推定を得るために、入力音声信号の1024点高速フーリエ変換が計算される(101)。マスキング閾値を決定するために、パワースペクトルが分析される(110)。従って、MPEG規格は、各サブバンドにおけるマスキング閾値と信号対マスク比(signal-to-mask ratio) (SMR) とを計算するために、音色信号成分(tonal signal component)、非音色信号(non-tonal signal component)、沈黙信号(silent signal component)の各々を計算に入れる(111)。

1) ビット割当ては反復的手順である(112、113)。

各々の反復ステップにおいて、量子化レベルの数と、従って、信号対量子化雑音比(signal-to-quantizing noise coderatio) (SNR) とが、マスク対量子化比(mask-to-quantizing ratio) (MNR)

MNR = SNR - SMR

が最小であるサブバンドにおいて、決められた増分だけ*

$$S(i) = \sum_{k=0}^{63} \cos \left[\frac{\pi}{64} (2i+1)(k-16) \right] y(k) \quad , i=0,1,\dots,31 \quad (1)$$

【0024】前式中では、上記規格で定義される通りに、ウインドウ演算と桁送り演算と加算演算とが行われた後に、64個のy(k)標本が、32個の入力音声標本x(i)から導き出される(302、303、304)。

【0025】入力y(k) (k=0, 1, 2, ..., 63)である時に、本発明のサブバンド分析フィルタに※

$$y'(k) = \begin{cases} y(k+16) & k=0,1,\dots,47 \\ -y(k-48) & k=48,49,\dots,63 \end{cases} \quad (2)$$

【0027】この時、

【0028】

$$S(i) = \sum_{k=0}^{63} \cos \left[\frac{\pi}{64} (2i+1)k \right] y'(k) \quad , i=0,1,\dots,31 \quad (3)$$

8

*増加させられる(105)。これと同時に、そのフレームを符号化するための他のビットが計算される(106)。上記反復手順が、残りのビットがある限り繰り返される。ビット割当て情報(107)が24ms毎に伝送され、従って、3.5kbit/sのデータ転送速度が必要とされる(108)。符号化されたビットストリームが、格納と加算処理等のために符号器によって出力される。

【0019】図4は、MPEG音声復号器でのレイヤI、IIにおける信号処理を示すフローチャートである。入力ビットストリームのビット割当て(200、201)が復号される(202)。その後で、様々な基準化因子も復号され(203)、標本が再量子化される(204)。符号化された信号が合成サブバンドフィルタで復号され(205)、その復号されたパルス符号変調信号が、更に別の処理及び/又はリアルタイム再生のために出力される(206、207)。

【0020】以下では本発明の好ましい実施例を詳細に説明する。

【0021】(MPEG符号器) 図5は、MPEG音声符号器のための(図3に参照番号102で示される通りの)分析サブバンドフィルタの動作を示すフローチャートである。この図は、主要なフィルタ動作、即ち、桁送り演算(データが64桁だけ桁送りされる)(301、302、303)、ウインドウ演算(出力標本を生じさせるために、桁送りされたデータが窓関数で乗算され、その後で加算される)(303)、復号された出力信号を生成する(306、307)ための行列演算(64×32行列ベクトル乗算が行われる)(305)とを示す。

【0022】MPEG音声符号化サブバンド分析フィルタにおける行列演算(305)は、次のように定義され、

【0023】

【数7】

※における行列演算は次の通りである。

1. y'(k)が次式の通りに定義される時に、データベクトルy'(k) (k=0, 1, 2, ..., 63)を形成し、

【0026】

【数8】

【数9】

(6)

9

10

【0029】2. $y'(k)$ が上記(2)で定義された *
通りである時に、次のようにデータベクトルを形成す *

$$\begin{aligned} y''(0) &= y'(0) \\ y''(k) &= y'(k) - y'(64-k) \quad k=1, \dots, 31 \end{aligned} \quad (4)$$

この時、

※【数10】

【0030】

※

$$S(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2i+1)k\right] y''(k), \quad i=0,1,\dots,31 \quad (5)$$

【0031】データシーケンス $x(i)$ ($i=1, \dots, N-1$) の離散余弦変換を $X(i)$ によつて表す。この場合に $X(i)$ は次の通りに定義され、

★【0032】

【数11】

$$X(i) = \frac{2}{N} e(i) \sum_{k=0}^{N-1} x(k) \cos\left[\frac{\pi}{2N}(2k+1)i\right], \quad i=0,1,\dots,N-1 \quad (6)$$

【0033】 $X(i)$ のIDCTは次式で表され、

☆【数12】

【0034】

☆

$$x(i) = \sum_{k=0}^{N-1} e(k) X(k) \cos\left[\frac{\pi}{2N}(2i+1)k\right], \quad i=0,1,\dots,N-1 \quad (7)$$

【0035】前式中で、 $i=0$ である場合に $e(i) = 1/\sqrt{2}$ であり、それ以外の場合には $e(i) = 1$ である。

◆ $S(i)$ は、 $y''(i)$ の32点IDCTから導き出される。

【0036】上記に示されているように、方程式(5)は 32×32 回の乗算累算を必要とする。上記方程式(5)を方程式(7)と比較すると、 $S(i)$ は(1つの基準因子の範囲内で) $y''(k)$ の32点逆離散余弦変換(「IDCT」)に対応する。従って、本発明のサブバンドフィルタでは、MPEG符号器行列演算の出力 ◆

20 【0037】上記の方程式(2)の定義と方程式(4)の定義とを組み合わせることによって、 $y''(k)$ と $y(k)$ との間の直接的関係が次のように導き出されることが可能である。

【0038】

【数13】

$$y''(k) = \begin{cases} y(16) & k=0 \\ y(k+16) + y(16-k) & k=1,2,\dots,16 \\ y(k+16) - y(80-k) & k=17,18,\dots,31 \end{cases} \quad (8)$$

【0039】(MPEG復号器) 図6は、MPEG音声復号器のための合成サブバンドフィルタの動作を示すフローチャートである。図4に示されているように、MPEG音声復号処理(レイヤI、II)は、同期化機能と、グループ化解除(degrouping)機能と、量子化解除(dequantization)機能と、正規化解除(denormalization)機能と、(図4に参照番号205で示されている通りの)サブバンド合成機能とを含む。図6では、32個のサブバンド標本が復号器(400、401)に入力され、桁送 *

*り演算(402)の後で、行列演算が行われる(403、404)。ウィンドウ演算(405)の後に、32個の復号された標本が計算され(406)、復元されたパルス符号変調信号が出力される(407、408)。

【0040】MPEG音声復号では、(図6に参照番号403で示される通りの)サブバンド合成フィルタにおける行列演算は、次式によって定義される。

【0041】

【数14】

$$V(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)(i+16)\right] S(k), \quad i=0,1,\dots,63 \quad (9)$$

【0042】上記の方程式(9)の面倒な計算は、 $32 \times 64 = 2048$ 回の乗算累積演算を必要とする。MPEG音声復号に関する性能分析は、行列演算が復号時間全体の約40%を占めることを示している。本発明のサブバンド合成フィルタでは、行列演算は、32点DCTを使用して方程式(9)を実行することによって復号を可能にする。

【0043】本発明のサブバンド合成フィルタにおける行列演算は次のように行われる。

【0044】方程式(9)から $V(i)$ が与えられる場合に、 $V'(i)$ は次のように定義される。

【0045】

【数15】

(7)

$$V'(i) = \begin{cases} V(i+48) & i=0,1,\dots,15 \\ V(i-16) & i=16,17,\dots,63 \end{cases} \quad (10)$$

【0046】上記方程式(9)と(10)から、次式が得られる。 * 【0047】 * 【数16】

$$V'(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)(i+64)\right] S(k), \quad i=0,1,\dots,15 \quad (11)$$

$$V'(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)i\right] S(k), \quad i=16,17,\dots,63 \quad (12)$$

$$V'(32) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{2}(2k+1)\right] S(k) = 0 \quad (13)$$

【0048】 V' が上記の通りに定義される場合に、 $i=1, 2, \dots, 16$ に関して $V'(32+i) = -V'(32-i)$ であり、 $i=17, 18, \dots, 31$ に関して $V'(32+i) = V'(32-i)$ である。 ※ $V''(i) = -V'(i)$ として定義され、 $i=1, 6, 17, \dots, 31$ に関して $V''(i) = V'(i)$ として定義される場合には、

【0049】 V'' が、 $i=0, 1, \dots, 15$ に関し ※

$$V''(i) = \sum_{k=0}^{31} \cos\left[\frac{\pi}{64}(2k+1)i\right] S(k), \quad i=0,1,\dots,31 \quad (14)$$

【0051】上記の方程式(6)と(14)とから、 $V''(i)$ は(1つの基準化因子の範囲内で) $S(i)$ の32点離散余弦変換に一致する。求められるべきべく ★ トル $V(i)$ は、 $V''(i)$ から次の通りに得られる。

$$\begin{aligned} V(i) &= V''(i+16), \quad V(32-i) = -V(i), \\ V(16) &= 0 \end{aligned} \quad (15)$$

であり、 $i=1 \sim 15$ の場合に、

$$\begin{aligned} V(48-i) &= -V''(i), \quad V(48+i) = -V''(i), \\ V(48) &= -V''(0) \end{aligned} \quad (16)$$

【0052】従って、MPEG音声復号における上記の行列演算方程式(9)の出力 V は、 S のDCT変換の後処理の後に得られる。

【0053】図7は、本発明によるMPEG音声復号器のための合成サブバンドフィルタにおけるDCTベースの行列演算のフローチャートである。従って、32個の標本が符号器に入力され(500、510)、 V'' が S の32点DCT変換から計算される(501、510)。この図の右側は、図の左側のステップ2からステップ4(502、503、504)が、 V が V'' から直接的に計算されることが可能であるように、1つのステップ(512)の形に組み合わせることが可能であることを示している。その後で、復号されたパルス符号変調信号が復号器によって出力される(505、513)。

【0054】本発明は、その一側面において、復号用にDCTを使用するか又は符号化用にIDCTを使用することによってサブバンドフィルタリング中の行列演算のより高速な実行を可能にすることによって、より効果的なMPEG音声符号化/復号化を可能にする。これに加

えて、本発明は、簡略化された余弦係数セットによって演算を行うことによって、システムの必要メモリを減少させる。

【0055】IDCTとDCTとの演算を実行するための手法が数多く知られている(K.R. Rao, P. Yip, Discrete Cosine Transform, Academic Press, 1990を参照されたい)。例えば、Leeの高速DCT演算は反復的であり、32点DCT(又はIDCT)に適用されることが可能であり、その結果として、合計で80回の乗算演算と209回の加算演算とが行われる(B.G. Lee, A New Algorithm To Compute The Discrete Cosine Transform, IEEE Trans. on ASSP, Vol. ASSP-32, No. 6, pp. 1243-1245, Dec. 1984)。このことは、従来技術で使用されるような面倒な計算に比較して、著しい計算上の節約を示している。

【0056】Leeのアルゴリズムは、効率的であり且つ理解と符号化/復号化とが容易なので、本発明の好ましい実施例で使用される。Leeのアルゴリズムに基づく32点高速DCTの実行を、次に説明する。

【0057】DCTデータシーケンス $x_i(i) = 0$,

(8)

13
1, ..., N-1をX(i)で表すと、N=32の場合に、高速DCTのための方程式が次式によって与えられる。

$$G(n) = X(2n), \quad n = 0, 1, \dots, 15 \quad (17)$$

$$H(n) = X(2n+1) + X(2n-1), \quad n = 0, 1, \dots, 15 \quad (18)$$

$$x(k) = g(k) + \frac{1}{2c_{64}^{2k+1}} h(k) \quad (19)$$

$$x(31-k) = g(k) - \frac{1}{2c_{64}^{2k+1}} h(k), \quad k = 0, 1, \dots, 15 \quad (20)$$

【0059】但し、
【0060】

※【数19】

※

$$c_{2N}^{(2k+1)n} = \cos(\pi(2k+1)n/2N) \quad (21)$$

【0061】図8A及び図8Bは、本発明によって実行されるLeeのアルゴリズムによる32点高速DCT演算に関するデータフローチャートである。方程式(17)から方程式(20)までが、(逆向きの矢印で)図8A及び図8Bに示されるフローチャートの第1の段階と最終段階とを形成する。両図における値H、G、g、hは、16点DCT演算と、8点DCT演算と、4点D★

★CT演算と、2点DCT演算とのために、同じ方程式を使用して反復的に計算される。MPEG符号化のためのIDCTを計算するために、図に示される処理方向が反転される。図8A及び図8Bでは、

【0062】

【数20】

$$C(2N, k) = 1/\left(2c_{2N}^k\right) \quad (22)$$

【0063】本明細書では好ましい実施例を参照して本発明を説明してきたが、本発明の思想と範囲とから逸脱することなしに、本明細書で説明した用途以外の用途に本発明が適用可能であることを、当業者は理解するだろう。従って、本発明は、添付クレームによってのみ限定される。

【0064】以上、本発明の実施例について詳述したが、以下、本発明を各実施態様に列挙する。

(1)

【0065】

【効果】以上のように、本発明を用いると、MPEG音声符号化/復号化処理における速度向上を実現することができる。MPEG復号では、本発明は、特別なハードウェアを必要とせず、音声復号における約30%の改善と、音声復号とビデオ復号とが組み合わされた復号全体における約17%の改善とをもたらす。

【図面の簡単な説明】

【図1】 MPEG音声符号器の略ブロック図である。

【図2】 MPEG音声復号器の略ブロック図である。

【図3】 MPEG音声符号器のレイヤIとレイヤIIとにおける信号処理を示すフローチャートである。

【図4】 MPEG音声復号器のレイヤIとレイヤIIとにおける信号処理を示すフローチャートである。

【図5】 MPEG音声符号器用の分析サブバンドフィ

ルタの動作を示すフローチャートである。

【図6】 MPEG音声復号器用の分析サブバンドフィルタの動作を示すフローチャートである。

30 【図7】 本発明による音声MPEG復号器のための合成サブバンドにおけるDCTベースの行列演算のフローチャートである。

【図8A】 本発明による32点高速DCT及びIDCT演算に関するデータフローチャートである。

【図8B】 本発明による32点高速DCT及びIDCT演算に関するデータフローチャートである。

【符号の説明】

10: MPEG音声符号器

11: 入力信号

12: サブバンドフィルタ

14: 量子化/符号化回路

13: 補助データ

17: 心理音響モデル

20: MPEG音声復号器

21: 符号化ビットストリーム

22: ビットストリームアンパック/復号回路

23: 補助データ

24: 復元回路

25: 逆マッピング回路

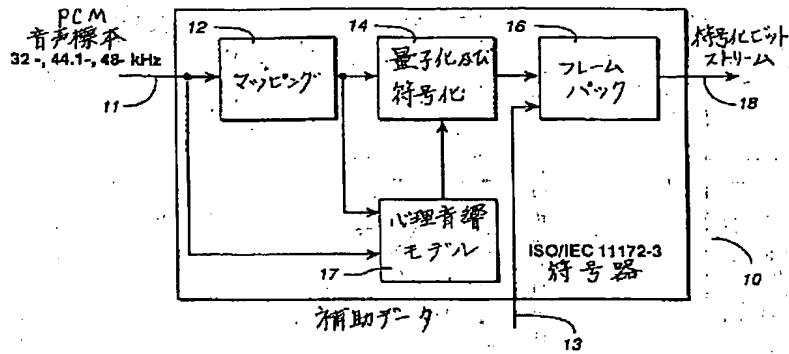
50 26: 出力信号

(9)

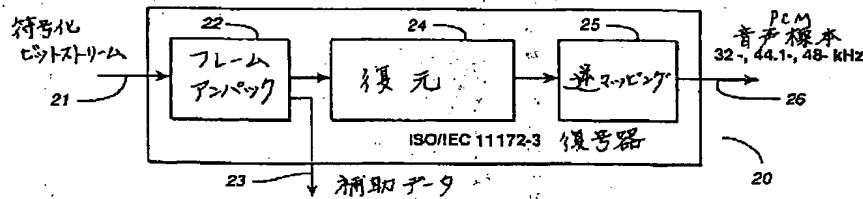
100、101、102、103、104、105、106、107、108、109、110、111、112

2、113:ブロック

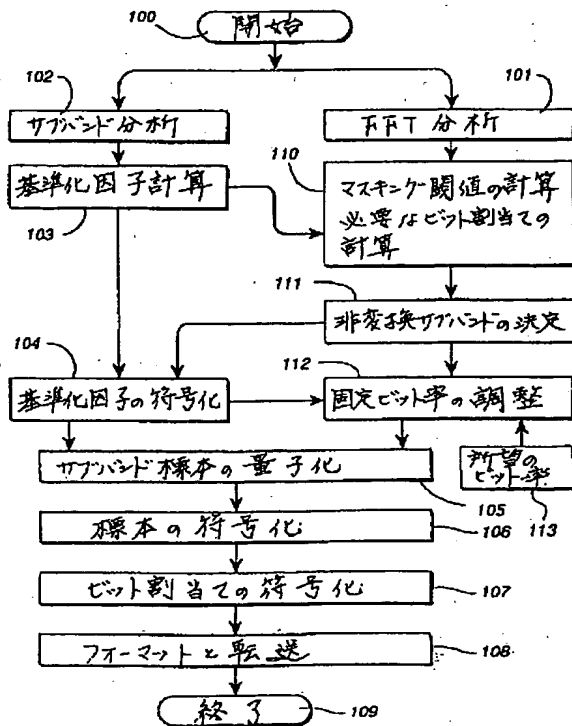
【図1】



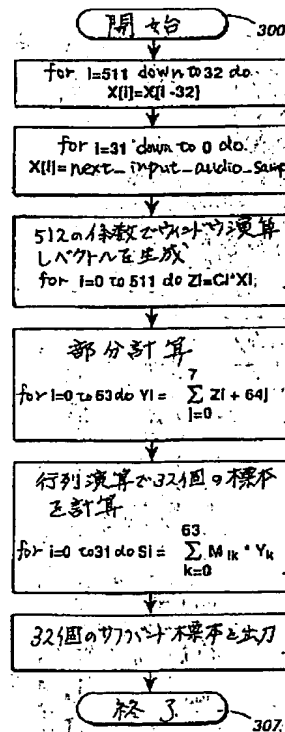
【図2】



【図3】

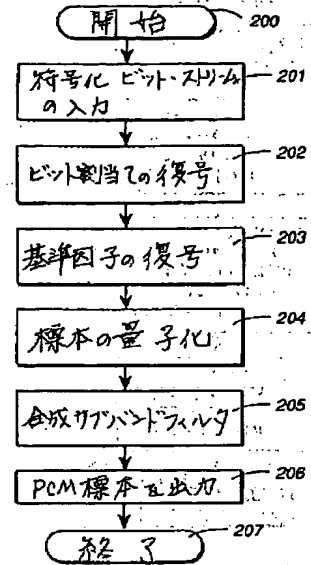


【図5】

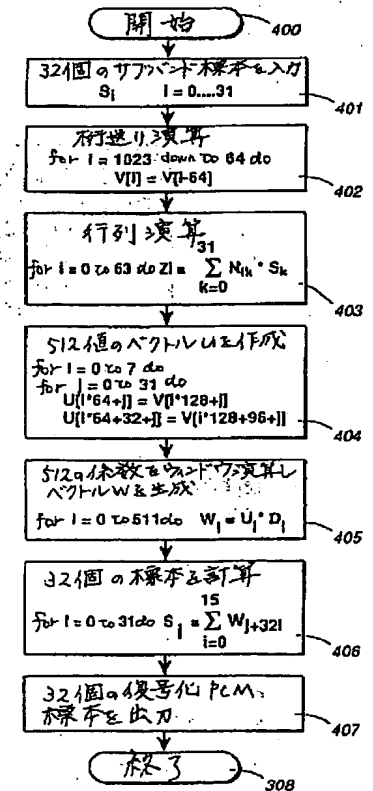


16

【図4】

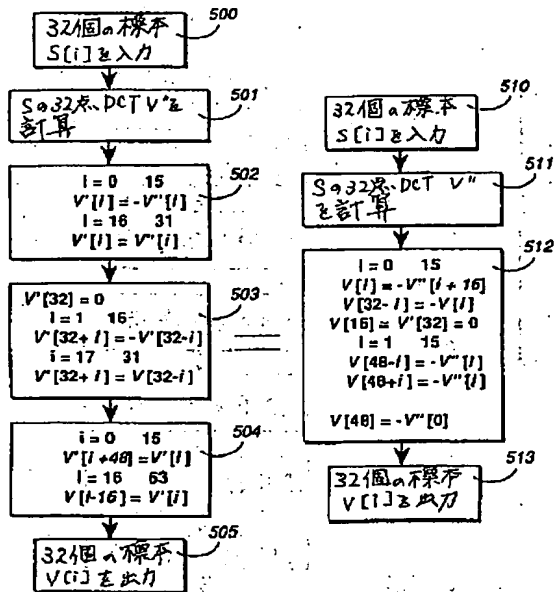


【図6】

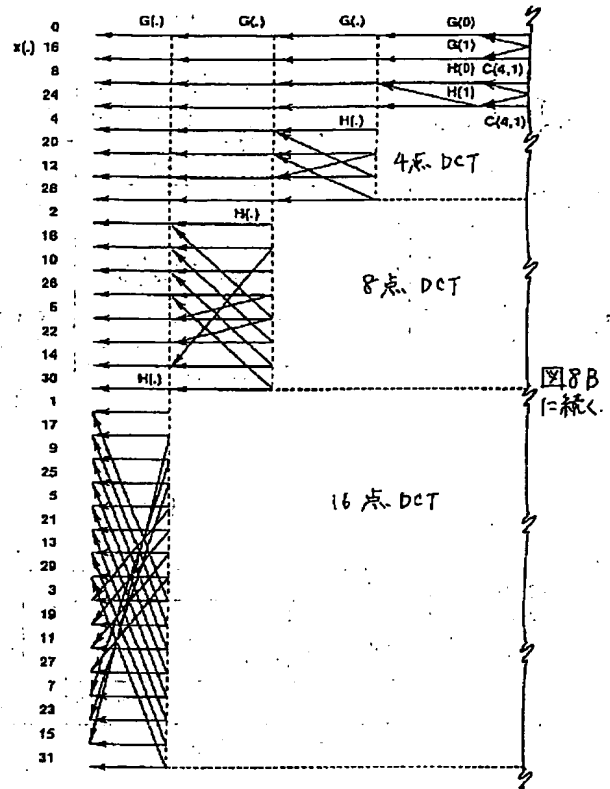


(10)

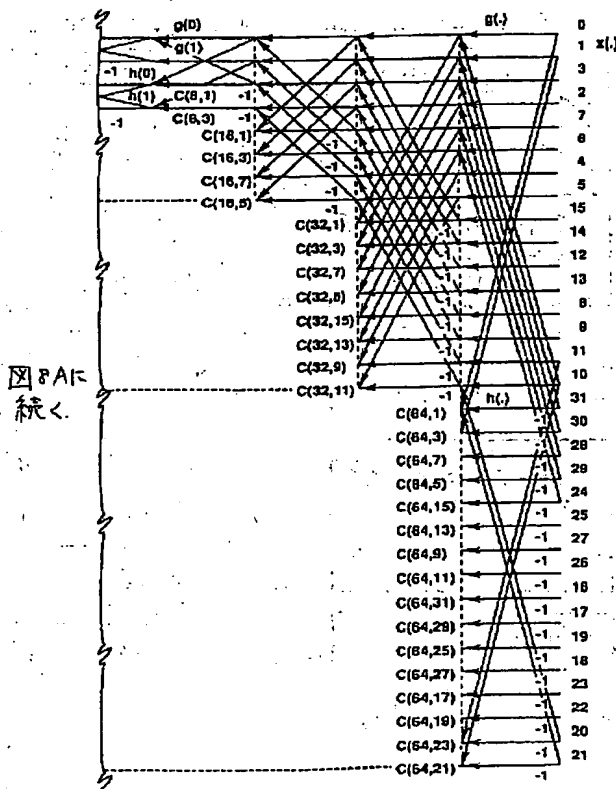
【図7】



【図8A】

図8B
に続く。

【図8B】

図8Aに
続く。

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.